



KDD-DDD 2017

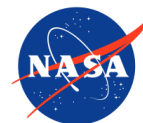
Measurement Context Extraction from Text

Discovering Opportunities and Gaps in Earth Science

Kyle Hundman, Chris A. Mattmann

Jet Propulsion Laboratory, California Institute of Technology

Copyright 2017 California Institute of Technology. U.S. Government sponsorship
acknowledged



Jet Propulsion Laboratory
California Institute of Technology

Motivation

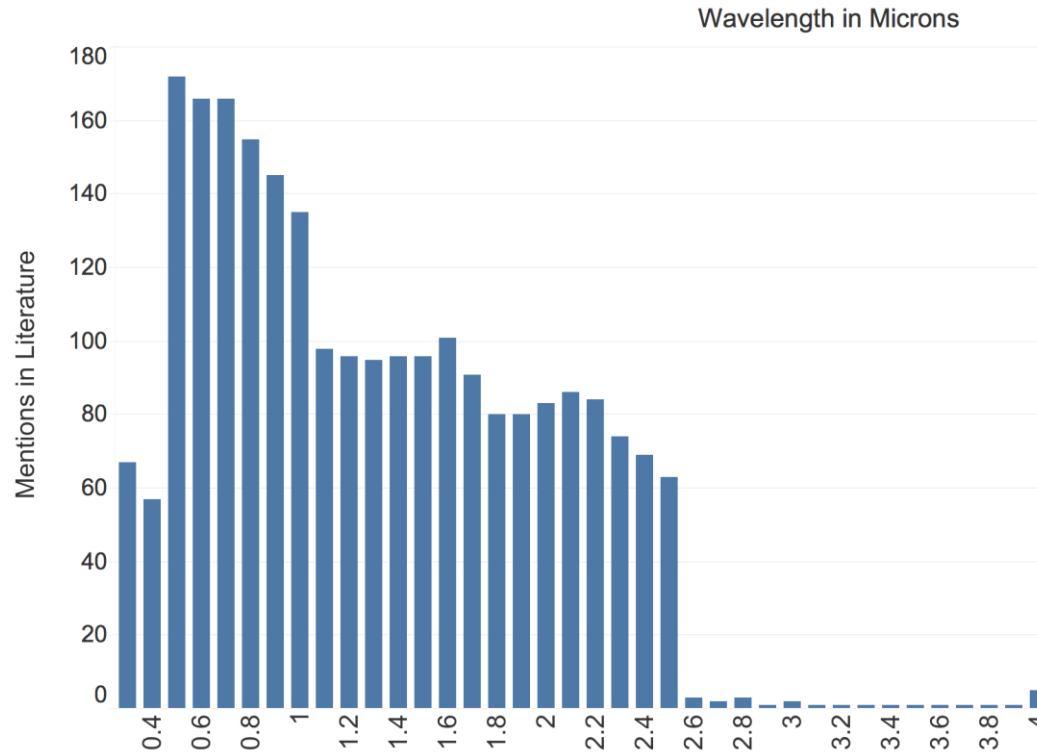
Beyond Measurements

- Scientific results and measurements often exist in text
 - Discussion = importance
- Measurements alone aren't valuable (ambiguity)
 - E.g. “50 m” in remote sensing domain could be...
 - Height
 - Swath
 - Length
 - Resolution
 - Need surrounding context to disambiguate

Motivation

Beyond Measurements

- Normalization and structuring enables deeper analysis



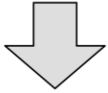
- Foundational tools are getting better



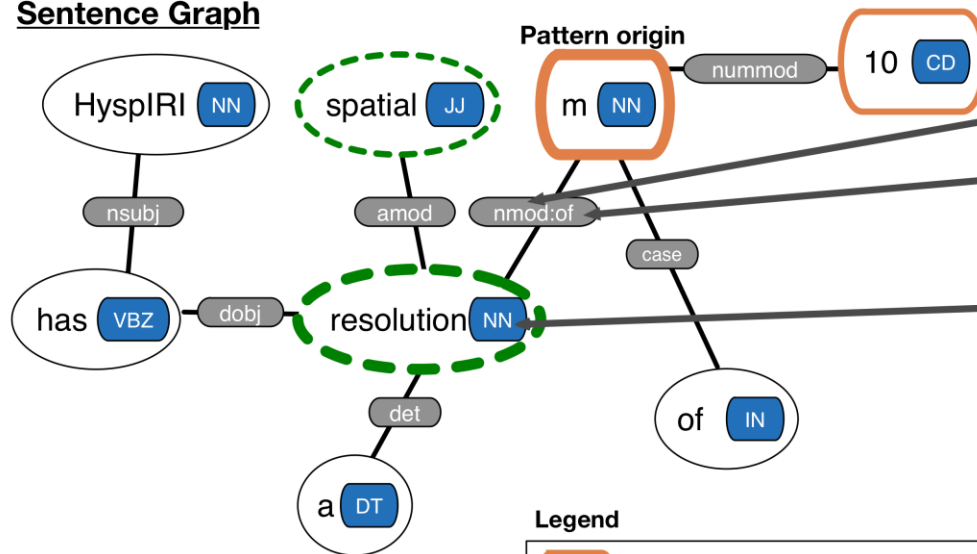
Parsey McParseface

Overview - Marve

Sentence: HyspIRI has a spatial resolution of 10 m.



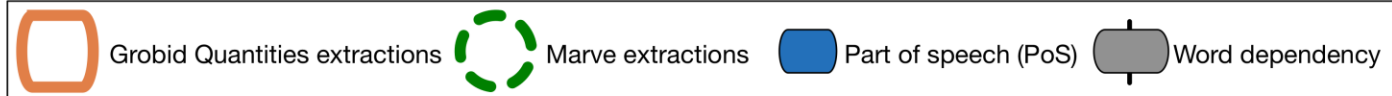
Sentence Graph



Marve Pattern Example

```
{  
  "nmod": {  
    "enhanced": true,  
    "of": {  
      "measurement_types": "space_between",  
      "pos_in": {  
        "NN": null  
      }  
    }  
  }  
}
```

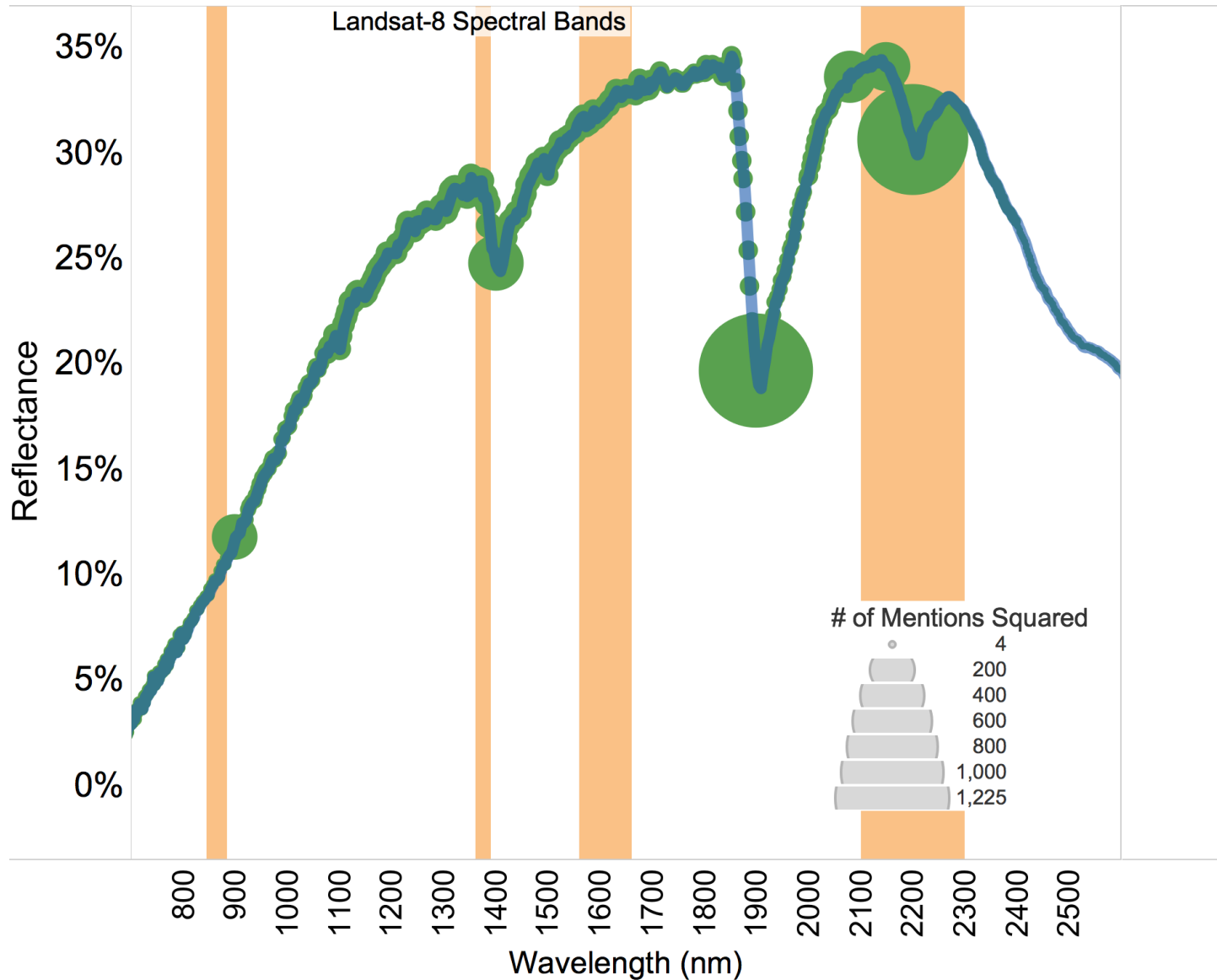
Legend



Measurement: 10 m

Related: spatial resolution

Use with HypsIRI



Future Work

- Refine patterns and test on more data
- Automatic grouping of measurements (still need post processing to do so now)
- **Contact**
- kyle.a.hundman@jpl.nasa.gov
- www.github.com/khundman/marve



Jet Propulsion Laboratory
California Institute of Technology

jpl.nasa.gov